

Indoor Tracking of Multiple Persons With a 77 GHz MIMO FMCW Radar

Nicolas Knudde ^{#1}, Baptist Vandersmissen [†], Karthick Parashar ^{*}, Ivo Couckuyt [#], Azarakhsh Jalalvand [†]
 André Bourdoux ^{*}, Wesley De Neve [†], Tom Dhaene [#]

[#]*Department of Information Technology, Ghent University–imec, Belgium*

¹*Nicolas.Knudde@UGent.be*

[†]*Department of Electronics and Information Systems, Ghent University–imec, Belgium*
^{*}*imec, Belgium*

Abstract—In this paper, we tackle the task of multi-target tracking of humans in an indoor setting using a low power 77 GHz MIMO CMOS radar. A drawback of such a high-resolution and low-power device is the higher sensitivity to noise, which makes the analysis of signals more challenging. Therefore, a pipeline is proposed to address both pre-processing of the radar signal and multi-target tracking. In the pre-processing phase, we focus on handling the low Signal-to-Noise Ratio (SNR) and eliminating so-called ghost targets. The tracking method we propose is based on Markov Chain Monte Carlo Data Association (MCMCDA), thus taking a combinatorial approach towards the task of tracking. The pipeline is tested on a number of real-world scenarios and shows promising results, overcoming the significant amount of noise associated with embedded radar devices.

I. INTRODUCTION

Radars are essential and widespread in numerous applications, including the navigation and tracking of different land, sea and air vehicles, as well as the automatic opening and closing of sliding doors. Moreover, radars are increasingly used in the indoor environments and are often incorporated as sensors in Internet-of-Things (IoT) applications. This is mainly thanks to advances in both hardware and software design, leading to higher resolution signals, miniaturized low-power radars and progressive signal processing techniques. They are among other devices applied in the surveillance domain, fall-detection of elderly [1] and gait analysis [2]. The advantages of using radar in comparison to using conventional visual sensors are the presence of explicit information such as range and velocity, while being light insensitive and privacy preserving. With such characteristics, radars are increasingly being deployed and analyzed in various use cases. In what follows, we briefly describe a few relevant references concerning multi-target tracking. Due to their high SNR, high-power radars have been used in many applications.

The Multiple Hypothesis Tracking (MHT) [3] method is one of the most frequently used techniques, utilizing the Maximum a Posteriori (MAP) estimate. Despite the effective performance of this method, it is computationally expensive due to the exponential growth of the number of hypotheses in time.

Recently, indoor through-the-wall tracking has also been considered, where the data exhibits multi-path interference [4].

Other works have focused on tuning a tracking filter for specific applications [5], [6].

The aim of this paper is to present an end-to-end system that enables the tracking of multiple persons in an indoor environment utilizing a low-power CMOS, Multiple Input Multiple Output (MIMO), Frequency Modulated Continuous Wave (FMCW) radar with a center frequency of 77 GHz and a bandwidth of approximately 1 GHz. We focus on the beam-forming mode, to retrieve both range and azimuth information, resulting in accurate estimates of the location of targets.

Our main contribution consists of an approach that is able to efficiently remove the high level of noise in the radar signals. As a result, we are able to provide features that are as clean as possible for the multiple object tracking algorithm. To that end, in Section II, we describe the application of the GoDec algorithm [7] to the radar data for modeling of the background scene, the combination of GoDec and Högbom CLEAN [8] for the identification of relevant targets and the use of a Bayesian technique [9] for track determination. Section III discusses the experimental setup and the results achieved by the proposed system. Finally, we provide brief conclusions in Section IV.

II. METHODOLOGY

In this section, we describe an overview of the pipeline developed to track multiple persons in a room using radar data. The flowchart of the pipeline and the corresponding steps are illustrated in Fig. 1.

Due to our aim of tracking multiple persons in an indoor environment, we need to distinguish the persons based on their spatial coordinates. This is not possible in the range–Doppler domain in which only the range and velocity data of the persons are available. In our approach multiple persons with the same range and the same speed, but with different angles (relative to the radar position) are distinguishable. On top, static persons can also be observed, whereas in the range–Doppler domain these would be absorbed by the clutter.

A. Pre-processing

The first stage of the pipeline consists of pre-processing the raw radar data in order to eliminate noise in the signal.

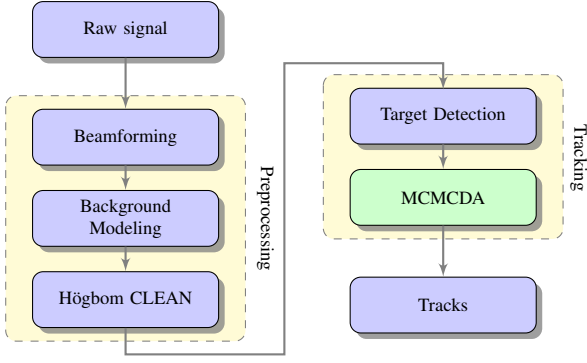


Fig. 1: The flowchart of the proposed approach for tracking multiple persons.

We focus on indoor scenarios that entail high amounts of clutter in the received signals. Subsequently, we describe three consecutive steps necessary to convert raw radar signals into clean data. However, it is worth to mention that the so-called “raw data” are actually the signals provided by the radar device after applying a built-in low-pass filter.

1) *Beamforming*: The processing of the raw radar data starts by applying the Fast Fourier Transforming (FFT) beamforming method [10]. The FFT transforms the input sequence of sampled data into a spectral representation of that data sequence. The output of this algorithm is the received power as a function of azimuth and range. Fig. 3a depicts the result of applying this algorithm to the raw radar data of a sample frame.

2) *Background Modeling*: Due to a potentially significant amount of reflecting objects in an indoor environment, the radar signal may contain multiple static targets that do not represent actual people. In order to effectively isolate targets from noise, ghost targets, sidelobes and static objects (such as walls and chairs), multiple pre-processing steps are performed and discussed in this and next subsection. It was chosen to replace the conventional Constant False Alarm Rate (CFAR) method for noise removal, often used for target detection.

As the first step, we employ the Go Decomposition technique (GoDec) [7] for the automatic modeling of the background scene in the radar data. This technique, which originally was developed for background modeling in video data, is based on the decomposition of the data X into three matrices:

$$X = L + S + G \quad \text{rank}(L) \leq r, \text{card}(S) \leq k \quad (1)$$

The low-rank approximation L (of rank less than r) represents the background, the sparse part S (of cardinality less than k) represents the part in the image that is changing and the matrix G represents random noise in the matrix. These are calculated by minimizing the Frobenius norm $\|X - L - S\|_F^2$ by using QR matrix factorizations. To the best of our knowledge, the GoDec algorithm has never been introduced to the microwave community.

Based on this technique we can accurately model the background and irrelevant objects of our environment. These

background data are subtracted from each frame to enable detection of relevant targets. The result of the background subtraction is shown in Fig. 3b.

3) *Högbom CLEAN*: An important aspect of the azimuth-range domain is the presence of sidelobes, which is due to the array characteristic and the beamforming algorithm. Sidelobes in itself can be confused with the real targets and should thus be removed. To that end, we make use of the Högbom CLEAN algorithm [8]. This is a technique to deconvolve the images, originally applied to radio astronomy images. It assumes that every target can be modeled as a point reflector. The array characteristic used is the Fourier transform of the virtual array (a sinc), where we made abstraction of the antenna characteristics. The algorithm iteratively subtracts the array characteristic centered at the maximum value in the image. However, as can be seen in Fig. 2c, it is not possible to filter out all sidelobes by using the Högbom CLEAN algorithm. Based on the specific situation and the number of iterations, some ghost targets remain visible in the processed frame.

B. Target tracking

Once the radar data have been processed, the track determination phase can be started. In principle, this phase consists of the following two steps:

1) *Target detection*: At each time step, the different targets present in the frame are detected. Initial candidate targets are detected based on the aggregated maximums found by the Högbom CLEAN algorithm (see Fig. 2c). Subsequently, the identified targets are filtered further based on their current power and a measure of their power in the previous and future k frames. More precisely, a target is identified as relevant if the product of its current normalized power and the Gaussian weighted power of this target in the arithmetic mean of a window of $2k+1$ frames exceeds some threshold. Fig. 3 shows an example of detecting two targets based on the power in the current frame and in the arithmetic mean of the according window of $2k+1$ frames.

2) *MCMCDA*: Tracking algorithms can be roughly divided into two categories: heuristic approaches and Bayesian approaches. Heuristic approaches do not work with an objective function, but for instance link the closest observations in space together [11].

Bayesian methods can work in multiple, more complex ways. One possibility is the use of a maximum a posteriori estimate of associations and the number of tracks at that estimate, taking into account previous measurements [3]. Another possibility is the use of a Bayes estimator to minimize the posterior expected value of some loss function [12].

In this work, we focus on using the Markov Chain Monte Carlo Data Association algorithm [9], which is a true approximation scheme for the optimal Bayesian filter. The method accounts for both missing measurements and false alarms. We use an adaptive version for online (real-time) tracking.

A Kalman filter is utilized to calculate the likelihood of a certain track. However, in the range-azimuth domain, the transformation becomes non-linear and an extended Kalman

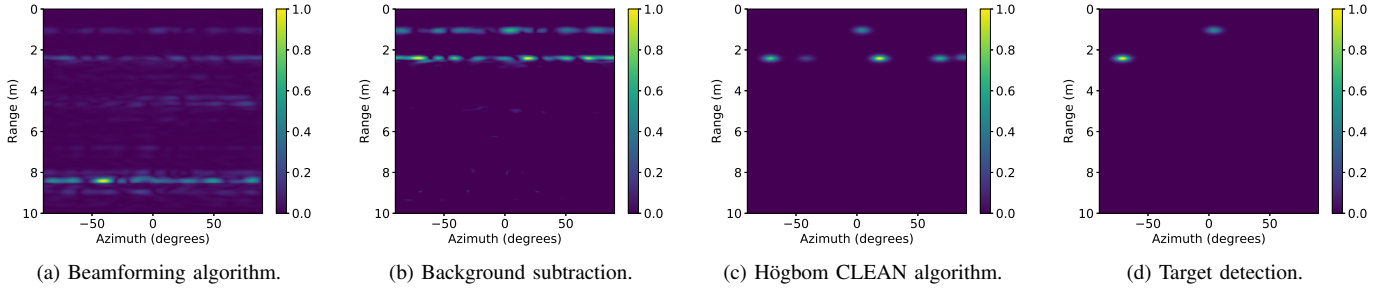


Fig. 2: The different processing steps, starting from raw radar data and ending with two identified persons.

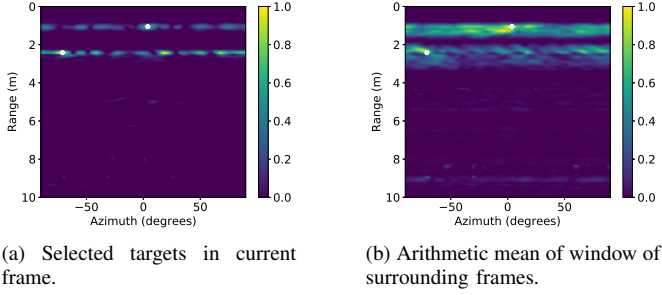


Fig. 3: Target detection based on power of target and window of frames.

filter has to be used. The state space vector is $\mathbf{x} = [x, \dot{x}, y, \dot{y}]^T$ and the applied process model is the constant velocity model:

$$\mathbf{x}_{k+1} = I_2 \otimes \begin{bmatrix} 1 & \delta \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k \quad (2)$$

with δ being the time between frames. Also, \mathbf{w}_k is the process noise modeled as having a Gaussian distribution with mean zero, $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$, and where \mathbf{Q}_k represents discrete white noise:

$$\mathbf{Q}_k = \sigma_Q^2 I_2 \otimes \begin{bmatrix} \frac{1}{4}\delta^4 & \frac{1}{2}\delta^3 \\ \frac{1}{2}\delta^3 & \delta^2 \end{bmatrix} \quad (3)$$

The measurement model simply is:

$$\begin{bmatrix} R \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{x^2 + y^2} \\ \arctan \frac{y}{x} \end{bmatrix} + \mathbf{v}_k \quad (4)$$

where the measurement noise \mathbf{v}_k has zero mean and a diagonal covariance matrix based on the resolution of the radar.

As shown in Fig. 4, the algorithm constructs different possible tracks by using a number of basic operations. After this, the posterior probability is evaluated as follows:

- The prior: evaluated using different parameters such as the number of track creations per unit time per unit volume, the number of false alarms per unit time per unit volume, the probability of a missing measurement and the probability that a target disappears.
- The likelihood: the likelihood probability per track is based on the marginal likelihood of the Kalman filter.

In the next step, a new track is constructed by one of the basic operations and the decision is made whether the new track is kept based on the Metropolis-Hastings sampling algorithm [13].

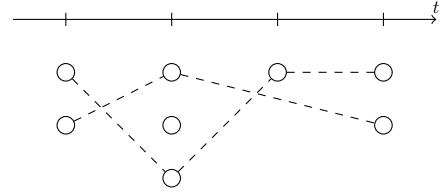


Fig. 4: Data Association schematic: the nodes represent the observations (including clutter) as a function of time t . The goal is to link observations to obtain tracks.

III. EXPERIMENTS AND RESULTS

The data is gathered with a low power MIMO FMCW radar. This radar consists of eight receiving and four transmitting antennas, making a virtual array of size 32. These are all used for the beamforming phase. The radar characteristics and processing parameters are given in Table I.

TABLE I: Radar characteristics and parameters

Parameter	Value
Output power	10 dBm
Horizontal beamwidth (RX)	76.5°
Vertical beamwidth (RX)	12.8°
Horizontal beamwidth (TX)	51°
Vertical beamwidth (TX)	13.2°
Bandwidth	1.1 GHz
Chirp duration	288 μ s
Repetition period	400 μ s
Dwell time	75 ms
Number of samples per upchirp	512
Sampling frequency	2 MHz
Range FFT size	2 ¹⁰
Azimuth FFT size	2 ⁹

The recording location (cf. Fig 6) has a construction based on a metal frame, augmenting the challenge of tracking, due to increased reflections from the boundaries of the room.

We recorded three proof-of-concept scenarios, ranked by increased level of difficulty: one person walking normally, two persons walking in parallel and two persons crossing. The tracking results for these scenarios are shown in Fig. 5.

The red data points on the plot show the detected targets by the target detection step. As can be seen, there is still a low number of false alarms present, as well as missing measurements. The full lines indicate the tracks determined by the MCMCDA algorithm in their Kalman filtered form.

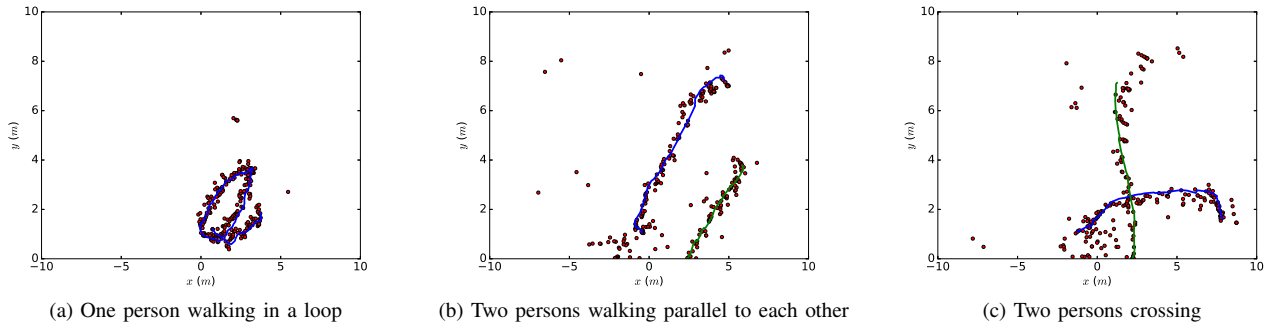


Fig. 5: Kalman filtered tracks for the three scenarios used. The dots represent the target detections and the outliers are false alarms

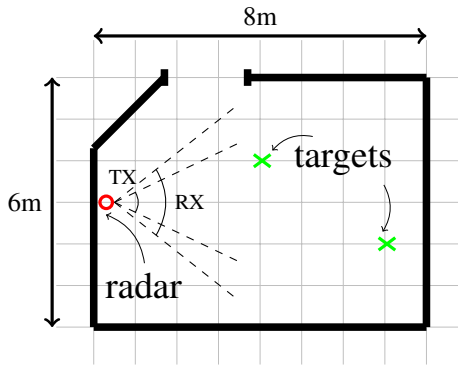


Fig. 6: Schematic diagram of the recording room.

In all three cases, the determined tracks follow the measurements and all false alarms are disregarded. The number of tracks is correctly estimated, and in the case where two people are crossing, the tracks follow the right person due to the constant velocity model of the Kalman filter.

IV. CONCLUSION

A data processing pipeline for indoor tracking of multiple persons using a low power, low SNR MIMO FMCW radar was presented. The pipeline was evaluated using three scenarios, showing that we are able to correctly track persons by identifying real targets and by coping with missing observations. In this work, we used the constant velocity model for the state space, which allows for the correct determination of tracks in case two targets are crossing. This study did not focus on the computational time necessary to generate the tracks; this will be investigated in more detail in future work.

ACKNOWLEDGMENT

Ivo Couckuyt is a post-doctoral research fellow of FWO-Flanders.

REFERENCES

- [1] B. Erol, M. Amin, F. Ahmad, and B. Boashash, "Radar fall detectors: a comparison," in *Proc. SPIE of Radar Sensor Technology XX*, vol. 9829, 2016.
- [2] D. Tahmouss and J. Silvius, "Radar micro-doppler for long range front-view gait recognition," in *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, Sept. 2009, pp. 1–6.
- [3] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, Dec. 1979.
- [4] J. Liu, G. Cui, Y. Jia, L. Kong, and X. Yang, "Sidewall Detection Using Multipath in Through-Wall Radar Moving Target Tracking," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 6, pp. 1372–1376, June 2015.
- [5] A. Vaidhyathan, S. M. Suhail, M. S. Babu, S. V. S. Rao, and V. S. Rao, "Design and evaluation of tracking filters for a special radar application," in *2012 International Conference on Communications, Devices and Intelligent Systems (CODIS)*, Dec. 2012, pp. 69–72.
- [6] C. Aydos, B. Hengst, and W. Uther, "Kalman filter process models for urban vehicle tracking," in *2009 12th International IEEE Conference on Intelligent Transportation Systems*, Oct. 2009, pp. 1–8.
- [7] T. Zhou and D. Tao, "GoDec: Randomized Low-rank & Sparse Matrix Decomposition in Noisy Case," in *Proceedings of the 28th International Conference on Machine Learning*, L. Getoor and T. Scheffer, Eds. New York, NY, USA: ACM, June 2011, pp. 33–40.
- [8] J. A. Hgbom, "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines," *Astronomy and Astrophysics Supplement Series*, vol. 15, p. 417, June 1974.
- [9] S. Oh, S. Russell, and S. Sastry, "Markov Chain Monte Carlo Data Association for Multi-Target Tracking," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 481–497, Mar. 2009.
- [10] B. D. V. Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, Apr. 1988.
- [11] X. R. Li and Y. Bar-Shalom, "Tracking in clutter with nearest neighbor filters: analysis and performance," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 3, pp. 995–1010, July 1996.
- [12] H. Tao, H. S. Sawhney, and R. Kumar, "Object tracking with Bayesian estimation of dynamic layer representations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 75–89, Jan. 2002.
- [13] W. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 97, no. 109, 1970.